How to Combine Survey Media (Web, Telephone, Face-to-Face): Lyon and Rhône-alps Case Study

Caroline Bayart, Patrick Bonnel

Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 11 (2015) 118 – 135

Transportation Research Procedia

www.elsevier.com/locate/procedia

10th International Conference on Transport Survey Methods

How to combine survey media (web, telephone, face-to-face): Lyon and Rhône-Alps case study

Caroline Bayart[a]*, Patrick Bonnel[b]

[a] *Laboratoire de Sciences Actuarielle et Financière, Université Lyon1, France, caroline.bayart@univ-lyon1.fr*
[b] *ENTPE, Laboratoire d'Economie des Transports, Lyon-Vaulx-en-Velin, France, patrick.bonnel@entpe.fr*

Abstract

We present the results of a web survey presented to non-respondents of a main telephone survey and compare declared mobility of both survey modes. After a description of the population who answered online, we summarise travel patterns and estimate a selection bias. Results are consistent with the hypothesis of an under declaration attributable to the web, as far as we know, that the risk of omitting trips concerns especially short trips and less constrained trip purposes. From comparison with a previous face-to-face and web survey we give some perspectives for future household travel surveys.

Peer-review under responsibility of International Steering Committee for Transport Survey Conferences ISCTSC

Keywords: Household travel surveys; web surveys; mixed modes surveys; data comparability; response rate; sampling coverage; design issues; selection bias.

1. Introduction

Survey response rates are decreasing across the world (Atrostic et al., 2001; Curtin et al., 2005). Household travel surveys follow the same tendency. Even if weighting procedures compensate for the incidence of non-response, it is always necessary to postulate that people with some socio-demographic characteristics who do not respond to a survey have the same behaviour as people with the same socio-demographic characteristics who respond. But evidence seems to indicate that it is not the case for travel (Richardson, 2000). Therefore, survey non-response might

* Corresponding author. Tel.: +33 472655427; fax: +33 472655412.
E-mail address: caroline.bayart@univ-lyon1.fr

Peer-review under responsibility of International Steering Committee for Transport Survey Conferences ISCTSC
doi:10.1016/j.trpro.2015.12.011

produce bias. Efforts to increase response rates for traditional surveys include improving the questionnaire, reducing respondent burden, and increasing reminders among others. Even if results are generally positive, it is in most cases not sufficient. According to Alsnih (2004), a way to increase the response rate and produce more reliable results is to propose a second media to answer, which initiated a project of a web survey in parallel to the household travel survey conducted by telephone in the Rhône-Alpes area. The main survey is conducted according to the CERTU methodology (CERTU, 1998): 12,417 completed interviews were conducted between October 2012 and March 2013. Households who refused to respond (about 27.5%) or were not contactable after a certain number of attempts (about 46%) were asked to respond by the web. Those households were informed by a letter of this option, followed by one reminder for only a small part of the sample due to budget constraints.

This new and interactive mode of data collection offers respondents the option to choose an appropriate time to complete the questionnaire, and does not require setting a phone appointment with the interviewer (Bourbonnais & Morency, 2012). However, the internet penetration rate is still uneven, and users' capabilities and equipment vary. If web surveys reduce the non-response rate, the implementation of a web survey raises specific problems, in terms of design and administration of the questionnaire. An online survey requires an auto-administrated questionnaire. It is important to work on the survey's attractiveness, simplification and technical feasibility to encourage potential respondents. The task is particularly complex for gathering individual trips. Lastly, if the launch of a web survey makes it possible to study behaviours little represented up to now (such as hyper-mobile households, with shifted schedules...), the question of data comparability remains (Stopher & Jones, 2003). The danger when databases are merged is that a sample selection bias will be created that will compromise the accuracy of explanatory models of travel behaviour.

This paper initially discusses web potential for household travel surveys, especially in a mixed modes framework (section 1). Then, some thoughts on the Rhône-Alpes area online questionnaire and the survey methodology are provided (section 2). We present the results of the Lyon web travel survey compared to the telephone survey (section 3), and characterise a selection bias (section 4). Lastly, we compare these results to the previous face-to-face and web surveys conducted in Lyon in 2006 (section 5) and give some perspectives for future household travel surveys (section 6).

2. Web potential for household travel surveys

Considering the democratisation of data processing and internet access, web surveys are promised as a fast method. Usually used in specific fields like marketing, it seems interesting to discuss the relevance of these new media for travel surveys, and determine under which conditions the use of the web can increase data quality, by taking into account the non-respondents to phone surveys. The main limit of the internet is the penetration rate, which is still too low in France for internet users to be considered representative of the French population. The idea is to ask some households to respond by web, while keeping the traditional method of data gathering by phone. This makes it possible to reach a large part of the reference population, and to study household travel patterns in-depth.

2.1. Why such a passion for web surveys?

Household travel surveys are generally long and expensive to implement. The objective is to survey a representative sample of households or individuals about their daily trips, using a quite complex questionnaire (Christensen, 2012). This kind of survey is usually managed in France face-to-face or by phone in medium French cities, with several household members being involved. But respondents are not always available to answer the survey and the non-response rate suggests many issues with this methodology. Variable costs are lower in a web survey than in a phone survey due to no staff. The use of the web media makes it possible to reduce the average cost of contact. On the other side, fixed costs might be higher, due to the development of the web questionnaire, but the design can easily be stored and re-used later. Thus, web media remains the least expensive, especially for large samples (Couper et al., 2001).

This mode of computer-assisted data gathering also allows an instantaneous checking of the provided answers (Dillman, 2007). To verify the validity of responses, dynamic controls occur throughout the questionnaire, and it is possible to send requests to the interviewees for clarification or correction in case of inconsistency. The quality of

collected data is better and, thanks to dynamic filters, respondents do not have to answer irrelevant questions and the clarity of the questionnaire reduces their level of frustration (Gunn, 2002). Moreover, online surveys are easy to implement due to no contact with respondents, and no sending of questionnaires. They allow real-time entry of data, without any physical effort to send back the questionnaire by post (Alsnih, 2004). Data gathering and data processing are thus faster than in other traditional modes. It is also possible to analyse respondent behaviour by collecting, for example, the time spent on each question, the number of partial or total stops, as well as the internet use pattern.

Time required to answer a survey is indeed limited and strongly constrained by the type of activities in which individuals are involved such as work, school, leisure, and various occupations in the residence. In general, self-administrated questionnaires are more flexible, since respondents can choose when to fill them out. Finding 20 minutes to answer a travel survey by phone in an overloaded schedule is sometimes a challenge. Internet connections are often personal, and web interviews can be carried out at home or at work, at an appropriate time. The assistance provided by the computer prevents interviewees from selecting questions, which often occurs in postal surveys. It is a dynamic mode, which recalls the respondent in case of inconsistency in the answers, informs them of the time remaining to finish the questionnaire, provides precise information before each question and allows some customisation. Lastly, the technology used makes it possible to fill out the questionnaire in part, with gathered data being stored in a database and recalled later thanks to a password.

2.2. Main limitations of web-based surveys

The potential of web surveys for collecting travel data seems to be high. However, this media presents some methodological limits, which should be considered. Usually, the authors distinguish four major sources of errors in web surveys (Groves & Lyberg, 2010; Alsnih, 2006): coverage error, sampling error, measurement error and non-response error.

To respond to an online questionnaire, households need to have a computer and an internet connection. Although multi-media equipment and internet penetration rate in France are growing (Table 1), they are not sufficient to permit the launch of a travel survey on the web only.

Table 1. France internet penetration rate.
Source: www.insee.fr (2013)

Years	2009	2010	2011	2012	2013
Penetration rate	71.6%	77.3%	77.8%	81.4%	81.9%

Moreover, we can differentiate internet users according to their equipment (navigator, connection type…) and their internet skills. According to the complexity of the questionnaire, some users will not be able to download pages, while others will have difficulties answering the questions with use of unrolling menus, question validation, and access to the following pages. So, even if the selected households have an internet connection, they may not be at ease with the use of these media. It is possible that not only the socio-demographic characteristics of the internet users differ from those of the whole population, but that they also vary according to the type of hardware owned (Alsnih, 2006). Thus, the sample generated by a web travel survey may be not representative of the target population, and gathered data cannot be generalised to the whole population by statistical inference (Bethlehem, 2010). Lastly, the presentation of the questions may vary according to the browser used and can modify the answers (Tourangeau et al., 2013). Because web surveys are self-administrated, the drop off rate in the course of the survey is high. Some technical issues are also specific to web surveys. Server unavailability or long data loading times can discourage certain respondents who abandon the questionnaire without the reason for this behaviour being known to the survey administrator (Dillman, 2001). It is not obvious that these non-responses, total or partial, are not related to the mobility behaviour (Bonnel, 2003). So, low response rates are a potential source of bias.

The administrator of a web survey must consider two other constraints: the cost and the control of the respondent. The time required to download web pages represents a cost for interviewees. It could be the time spent in front of the screen, or the connection length, for those who do not take advantage of an unlimited package. If variable costs

of questionnaire management and data gathering are marginal, a web survey generates important fixed costs, as the development of an online questionnaire requires time and specific skills. In addition, it is difficult to control "who" is in front of the computer. This limitation, applicable to self-administrated questionnaires as the virtual interviewer cannot check the identity of the surveyed person, affects the relevance of collected data. The respondent cannot obtain precision on concepts, and can easily omit some answers or "click" without thinking of the meaning of the questions. To conclude, data confidentiality remains an issue (Thériault et al., 2012). Although some internet users are familiar with contact customisation (Yun & Trumbo, 2000), the intrusion of data processing in private life and proliferation of computer viruses are disturbing. It seems better to ensure data confidentiality by using protected servers, and setting up fire walls to guarantee file safety.

2.3. Combining survey modes to improve data quality

It is still not recommended to conduct only a web survey if the aim is to gather representative data from the general population (Bonnel, 2003), but it is important to explore the combination of several survey modes (Couper & Bosnjak, 2010; de Leeuw, 2005). In our study, the combination of web and telephone seems interesting, since it could be a way to reduce sampling error: households who are presented with a web questionnaire have been recruited before by a probabilistic method (even if only those who have an internet connection are able to respond). Moreover, introducing the web into a large travel survey contributes to diffusing this media, and is a way to convert households formerly faithful to the traditional survey modes. Lastly, by proposing a second mode to gather personal information, it is possible to increase data quality by recruiting more statistical units. Respondents appreciate having several media available to answer, and each survey mode does not target the same segment of the target population. Some people can only be contacted by one type of media (such as those often out of home), or are especially willing to answer if this media is available.

However, web and telephone have very different characteristics, such as media used, design and length of the questionnaire, which directly impact their performance. During a phone interview, communication is primarily auditive, whereas the web uses visual language such as numerical texts, graphs and information which appear to the internet users (Couper et al., 2001). As using multi-media allows the combination of several languages, the response rate and the quality of computed information could vary (Dillman et al., 1998). In addition, the presence of an interviewer can generate a bias. For example, the respondent is likely to position his speech according to the positive impact he hopes to produce on his interlocutor ("social compliance"). Although this phenomenon especially occurs in opinion surveys, it can also be noticed in travel surveys, as transport is a sensitive field. For instance, in a survey being financed by a public transport authority, individuals may favour public transport. This bias is strongly reduced in the case of a web survey, since the interaction between the interviewer and the respondent is virtual. Moreover, when a list of items is read by the interviewer, respondents often select the last one in the "recency effect". But when the questionnaire is auto-administrated, respondents prefer the first item in a "primacy effect". A mixed modes survey using the web and telephone interviews generates comparability issues (Dillman & Browker, 2001). Lastly, mode effects could damage data quality (Gunn, 2002). Some cultural differences are related to the media chosen to participate in a survey (Morris, 2003). The households who respond by internet may have different travel patterns from those who prefer a phone interview (Bayart & Bonnel, 2012; Kagerbauer et al., 2012). These differences should be considered in data processing, in order to produce representative results (Bayart et al., 2009).

3. An application: the Rhône-Alpes web survey

In order to include non-reachable households or those who do not want to complete an interview by phone, the Laboratoire d'Economie des Transports proposed to implement a travel survey on the internet. The design of the web questionnaire is sensitive, since we need to take into account the constraints of the standard version developed by CERTU (1998), but also consider web-specific issues. If the web can increase the response rate and improve data quality, it is important not to increase respondent burden by setting up too complex a design (Bonnel, 2003). Technical considerations should balance the user-friendliness of the questionnaire.

3.1. An adaptation of the CERTU standard questionnaire

The CERTU standard questionnaire used in France for household travel surveys by phone is organised in four parts to gather successively: characteristics of the household and housing, socio-economic characteristics of each household member, daily trips of one or two random household members over 11 years old (carried out on a reference day), and opinions of a household member over 16 years old. Even if the questionnaire is relatively short to fill in (20 minutes on average), and easy to put on the web, since telephone interviews are assisted by computer (CATI), we had to make some simplifications. Some opinion questions were removed and others were adapted. For example, geographical data was coded using maps for the telephone survey but through interactive geolocalisation for the web survey.

In the phone survey, one or two household members could be surveyed, according to the household size. Technically, it would have been possible to specify this constraint in the letter sent to introduce the web survey. But, this objective seems difficult to achieve, since the main asset of this media is its flexibility. So, only one household member (over 11 years old) is asked to fill in the questionnaire. Thus, the standard questionnaire has been simplified, since people doing web surveys are often in a hurry, and do not want to spend too much time responding.

3.2. Data collection

The Rhône-Alpes web questionnaire is structured in four parts, or "blocks". We distinguish questions relating to the household, the respondent, daily trips data, and complementary questions about the willingness to participate in future travel survey and level of income. This order differs slightly from the phone interview, but in a self-administrated questionnaire personal information is sensitive to gather. To not worry the interviewee, we chose to begin the questionnaire with simple questions about car ownership, usual mobility behaviour, and the main socio-demographic characteristics of the household members. It is possible to quickly establish a link with travel patterns and to return to questions relating to the computer equipment and the annual income of the household at the end of the questionnaire.

Individuals who choose to answer a web survey understand there is some rigidity in web surveys, but no interviewer is present to motivate them. To reduce the abandonment risk we left a loophole ("Don't know" or "Other, to specify") as often as possible, which is useful for interviewees who do not want to answer specific questions. Moreover, gathering trip data requires time and great memory effort from respondents. To reduce the burden, intermediate validations are used: after each block "household" and "individual", after each trip, and at the end of the questionnaire. The idea is to store the answers, which could be recalled later in case of abandonment. Lastly, many filters occur within the questionnaire. They are more numerous than in a phone survey, since there is no interviewer to guide the respondent. The aim is to avoid redundant questions, which would lengthen the duration of the questionnaire and limit its credibility.

3.3. Information relating to the household and the respondent

Despite only one household member filling out the questionnaire, it is important to know the characteristics of the other household members, in order to understand household mobility patterns. The first question asks for the age, the sex and the occupation of all household members, based on information provided by the respondent. First names are also required, in order to customise the questionnaire and establish a link with personal questions. The gathering of home address must be accurate. As this variable is extremely important to analyse individual mobility, a detailed frame is proposed and online geocoding with Google Map and Open Street View assists the respondent. After considering the number of cars held by the household, we ask about possession of a driving licence. This variable is strongly explanatory for trips in a private car. Then, there are questions concerning the level of study, the type of occupied employment, and the accurate professional address. Several questions on the frequency of use of transport modes (walking, public transport, private car, bicycle…) for daily trips supplement the individual questionnaire. We added questions concerning the opportunity to answer an internet travel survey again or to participate in a GPS study. As these questions are not directly related to mobility, and can be seen as an intrusion in the private sphere, we put them at the end of the questionnaire. The income classes are similar to those used in the

phone survey. We assume that the "Don't know" answer will be selected, as people are not used to revealing information related to their purchasing power in France.

3.4. Information related to daily trips

In front of a computer, respondents do not consider a survey in the same way as when faced with an interviewer (Dillman & Bowker, 2001), since non-verbal aspects of the questionnaire such as graphics, logos, and interactivity can influence the answers (Couper, 2008). This result is particularly important for travel surveys, where individual data remains difficult to gather. Generally, the concept of "trip" is not easily understandable by respondents. To avoid measurement bias due to different interpretations of this word, interviewers usually give some information by phone. However, web surveys do not allow this interactivity, and the understanding of "trip" by internet users remains difficult to control. As this experiment was innovative, we needed to specify at the beginning of the questionnaire the objective of the approach and the instructions required to answer the questions. The objectives were to enhance the value of the answers, to insist on the accuracy of provided data and to explain how to obtain some assistance. Then, a fictitious example was given to illustrate the recommendations. In spite of the risk of suggesting answers, it allowed respondents to understand how to answer the questionnaire, did not take much space (one page), and exploited web multi-media potential.

We chose to ask questions about activities on the reference day, in order to collect the maximum information in a natural way for the respondent, so the memory effort was less important (Figure 1). The respondent has to notify where she/he was at the beginning of the reference period, as well as the activity carried out there. Then, questions follow the sequence below, until the time of departure of the last trip at the end of the reference period.

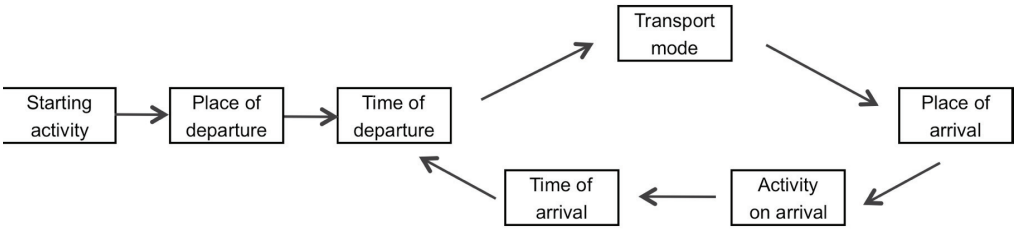

Figure 1. Structure of trip data collection.
Source: web questionnaire developed for the Rhône-Alpes area household travel survey (2012)

As anticipating respondents' connection date is impossible, a reference day is used to gather trip information. It corresponds to the last weekday before the day of first connection, and is automatically displayed. This method induces two principal issues: if a lot of people are connecting during the weekend, Friday trips will be over represented. Moreover, the memory effort will be more important for internet users who answer on Sunday or Monday. But it is not obvious to remember all trips carried out the day before in a chronological order. Some respondents probably wish to modify, add or remove data. At the end of the block "trips", a summary table of answers provided guarantees a reliable display of the data, whatever the type of hardware used and the number of trips entered. Respondents have to check this information, and can modify it or insert a new trip. As suppression is an irreversible action, a message appeared to avoid handling errors, which can lead to abandonment. Lastly, there were some coherence tests. The objective is to check the structure of the response and to detect possible contradictions as the respondent fills out the questionnaire.

It is also important to consider the gathering of trip origin-destination data. In the usual phone survey, all the big city areas in Rhône-Alpes are divided into zones. However, it is difficult to set up this process online, and to ask internet users to choose the appropriate zone. The abandonment of zoning enhances the collection of accurate data,

with precise addresses being entered by the respondents and geocoded by Google Map and Open Street View tools, and later by the administrator if there is no address linked to the information given by respondents.

4. Results of the fieldwork

The implementation of an online questionnaire is useful not only for the survey administrator, but also for all the households who are familiar with the internet. But, online respondents are more frequently present in specific categories such as young person, male, with a high level of education and income, and large household. If the socio-economic characteristics of households who answer on the internet differ from those of the target population, their mobility habits may not be the same (Bethlehem, 2010). First, we present the general results of the survey, in terms of response rate and response behaviour. Then, we describe the population who answered online, and summarise travel patterns.

4.1. Response rate and answer behaviour

Phone interviews took place between October 2012 and March 2013. By the end of the period, 10,695 households answered the survey and 12,417 interviews were conducted. As 40,517 households had been contacted, the response rate was around 26.5%. Households who did not want to respond by telephone received a letter which asked them to fill out the questionnaire by internet. A similar letter was also sent to the households not contactable by phone after at least 10 attempts. These letters especially highlighted the new media available to participate in the survey, which is more flexible. The letters were sent by mail because the electronic address of the selected households was not known. The letter mentioned the address of the website where the questionnaire was available, and a moderate response time of 15 minutes, as well as a personal login and password. Six waves of letters were carried out, in order to limit the time between the refusal and the new request.

As of 28 April 2013, around 2,030 individuals reached the site to try to answer the survey. But only 1,626 gave a complete and usable answer, representing an answer rate of 7%. These results seem low, compared to the telephone survey, but are promising if we consider that households have refused or were not able to participate in the phone interview first.

The average time required to answer the web questionnaire is close to the time noticed during the previous web survey in Lyon (19 minutes 20 seconds vs 18 minutes 36 seconds) and standard telephone travel surveys. Disparities are however important (between 4 minutes and more than one hour). The longest survey part is the report of daily trips (more than 10 minutes on average). This is due to the number and the precision of the questions, which require an important memory effort for addresses of visited locations, and departure and arrival times. For the other parts of the questionnaire, response times were shorter, around 3 minutes on average for each part. “Household” and “individual” blocks included more usual questions for interviewees, and often had drop-down lists for the answers. This is simple for the internet respondents, who minimise their connection time. During the week, internet users connected from 8:00 am till 11:00 pm, with three different peaks: 10:00 am, 2:00 pm and 6:00 pm. It is likely that some connections were made in the workplace. On the weekend, internet users connected from 9:00 am till 10:00 pm, with three different peaks: 11:00 am, 5:00 pm and 9:00 pm.

4.2. Who are web respondents?

Comparison of data from the standard household survey and web survey allows us to describe the profile of web respondents. While differences in terms of residential location are not significant, other socio-economic characteristics stand out clearly (Table 2).

The households who answered the web survey have more members (2.36 persons vs 2.22). The proportion of single person households is lower (25% vs 32%). Internet households are more motorised: 55% possess at least two cars compared to only 45% in the standard survey. The proportion of households without any car is lower: 6% vs 11%. The conclusion is the same for the number of vehicles by persons old enough to drive. This greater car access is partially explained by the size of the household and the socio-economic characteristics of the web respondents.

Web respondents have a higher monthly income by number of persons of the household (1,400 euros vs 1,270 euros on average). This difference can be explained by the fact that the penetration of internet is still uneven, and requires higher income due to the cost of the investment in computer hardware, internet subscription and higher level of study. Moreover, those who have a connection at the workplace have generally superior occupations, with a higher level of income. Finally, the proportion of households refusing to give their annual income level is less important on the web (one household in seven vs one household in five). Respondents are reluctant to communicate personal information, especially when it deals with their income. It seems more difficult to declare income by phone to an interviewer than on the web, which is by definition auto-administered, but it is also possible that those who accept to respond by internet are then less reluctant to give personal information.

Individual data analysis allows us to refine this comparison. The proportion of men who answered on the web is slightly higher than observed in the phone survey (52% vs 44%). On the other hand, the proportion of active persons (between 30 and 59 years) who answered on the web is higher (two-thirds vs 50% in the standard survey). They belong to high social and occupational groups (47% are executives or have a superior intellectual profession vs 26%) and have a high level of studies (about 55% made studies vs only 37% in the standard survey). These two variables are linked, and can be explained partially by the person who chose to fill out the questionnaire (the household leader or his spouse). Web respondents more frequently have a driving licence (96% vs 84%), with the oldest and youngest people being under-represented in the web survey. But the percentage who work or study in the town centre does not vary according to the survey mode (around 33%). The conclusion is the same for housing (52% live in a private house).

The web population seems to have particular characteristics. Active persons (64% vs 45%) may be less available to answer a long questionnaire by telephone. They are probably more sensitive to the survey subject and familiar with the internet, and agreed to answer the online questionnaire. But they are less likely to answer a travel survey by web again (72% vs 87%) or engage in a GPS survey (23% vs 29%).

Table 2. Socio-demographic characteristics of the individuals by survey.
Source: web and telephone travel survey - Rhône-Alpes area 2012

Variables	Levels	Phone survey	Web survey
Number of persons	1	32%	25%
	2	38%	41%
	3	12%	14%
	4	12%	14%
	5+	6%	6%
Gender	Male	44%	52%
	Female	56%	48%
Driving licence	yes	84%	96%
	no	16%	4%
Age	< 30 years old	19%	9.5%
	30 to 59 years old	49%	65.5%
	> 59 years old	32%	25%
Occupation	Executive	26%	47.5%
	Employee	44%	33.5%
	Other status	15%	17.5%
	No job	15%	1.5%
Next web survey	yes	87%	72%
	no	13%	28%
Next GPS survey	yes	29%	23%
	no	71%	77%
Educational level	Higher	37%	55.5%
	Not higher	63%	44.5%

4.3. Mobility pattern

It has been shown that new technology, particularly the internet, has a major impact on activity patterns, and hence travel practices (Mokhtarian et al., 2006). It is therefore very likely that the socio-economic differences between the two groups affect the travel practices of respondents (Ettema et al., 2007).

We have reweighted the face-to-face sample in order to have a similar distribution for the main socio-economic variables. The weighting process gives weight to each individual in the sample in order to obtain a population similar to the target population regarding some socio-economic variables. We preferred to reweight the phone sample rather than the web-based sample because the sample size of the web survey is smaller. It avoids giving high weight to some individuals belonging to a small group which can bias some average indicators. Margin calibration methods are commonly used to weight data. They perform the calculation of weight for each observation in order to respect known variable distributions, or in our case the web-based sample distribution for some socio-economic variables (Table 2). We have used Calc function available for R software. This method has been developed by the French National Institute for Statistics and Economic Analysis (Institut National de la Statistique et des Études Économiques). After calibration, profile differences between the web-based and telephone sample are considerably reduced. Seven socio-economic variables have been used: age, gender, number of persons in household, educational level, driving licence, occupation, future participation in a travel web survey and future participation in a GPS survey. Distributions are also closer for other variables which have not been taken into account because correlations between variables are often important such as income, car ownership, and economically active or not.

Having comparable samples on the basis of socio-economic characteristics, we analyse the mobility pattern of these respondents. Web respondents declare fewer daily trips than people interviewed by phone (3.09 vs 3.67) which contradicts some studies which have found that using new technology generates more activities outside the home and hence more trips (Wang & Law, 2007). This gap is explained by a higher number of immobiles (17% vs 11%) and a lower mobility of mobiles in the web survey (3.73 vs 4.05 daily trips). It might be at least partly due to the fact that by phone, an interviewer can make reminders and control the relevance of the trip pattern. In contrast, by web, the respondent is alone in front of the computer, and can easily omit “small” trips or return to home in order to voluntarily reduce the survey length.

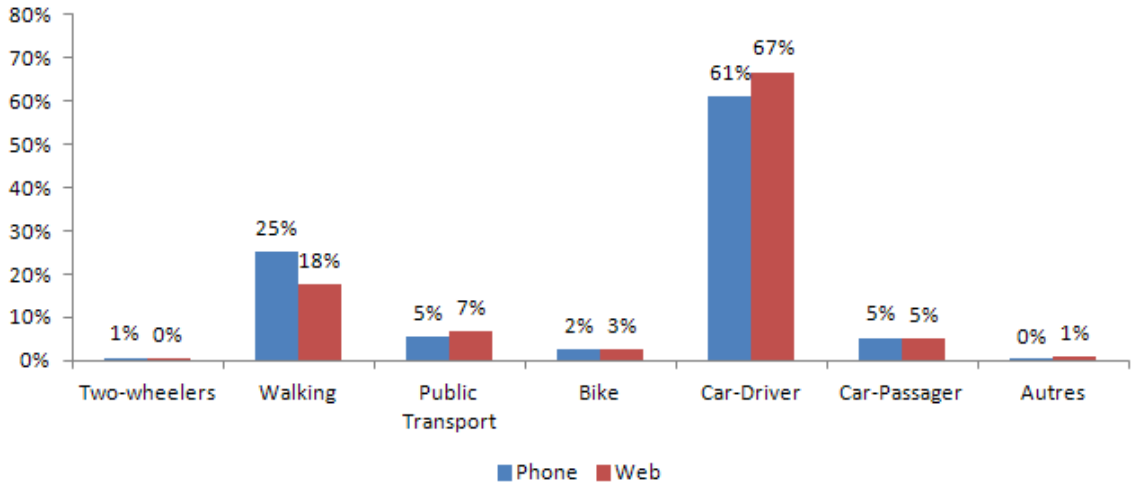

Figure 2. Distribution of daily trips by transport mode and by survey.
Source: Households web and telephone travel survey - Rhône-Alpes area 2012

Web respondents drive a car more often (67% vs 61%), and walk less (18% vs 25%) (Figure 2). But if we add the data for all motorised modes, the mobility of web respondents is greater than mobility of phone respondents, especially if we exclude those who did not move. It would thus seem that the web survey “correctly” collects

motorised trips of mobile persons, but leads to a lower declaration of walking trips, which are also "small" trips. The distribution of daily trips by purpose at the destination and by survey is the same in both surveys even if shopping and services trips seem to be under-represented in the web survey (Figure 3).

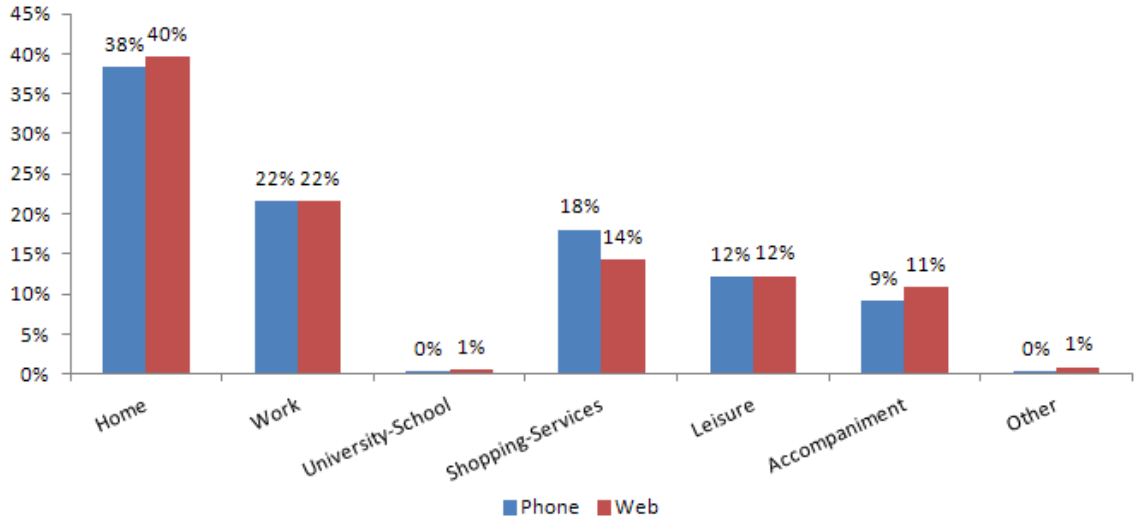

Figure 3. Distribution of daily trips by purpose at destination and by survey.
Source: Household web and telephone travel survey - Rhône-Alpes area 2012

Web respondents cover longer distances on average (9.3 km vs 6.4 km) and their trips have a longer duration (25.1 minutes vs 19.7 minutes). In the web survey, we noticed a strong proportion of trips made in a private car from home to work, but few walking trips which explains why declared distances and duration of trips in the area of study are higher on average. Lastly, web respondents seem to return home later in the evening (Figure 4).

In the following subsection, we seek to partially neutralise the socio-economic characteristics of the respondents. The issues which then arise are the comparability of the data and the methodologies which can be used to separate those differences in responses which are due to the methodology or to real differences in the behaviour of each group.

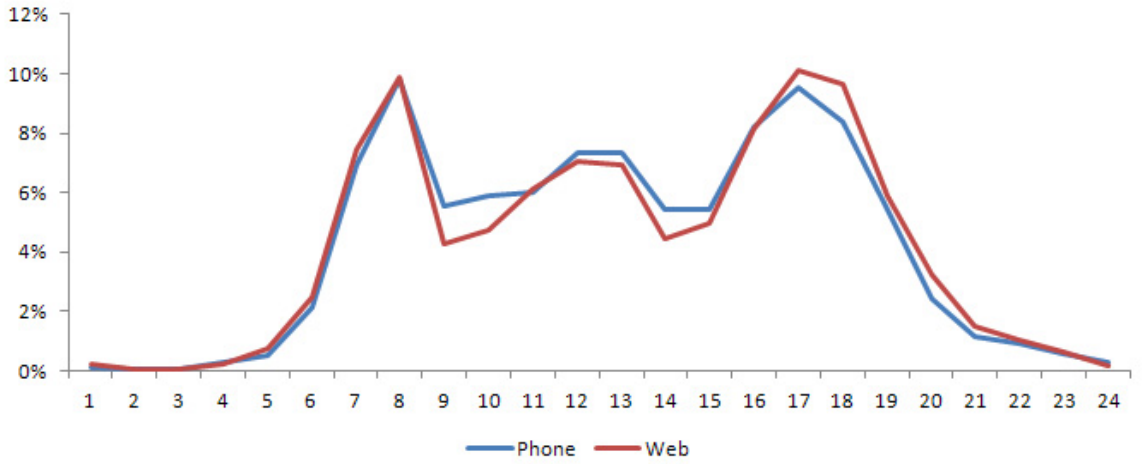

Figure 4. Average distribution of daily trips by departure time.
Source: Household web and telephone travel survey - Rhône-Alpes area 2012

5. Sample selection bias

It has been known since the 1950s that estimating an equation on the basis of a subsample selected from the population may result in biases (Roy, 1951). However, the first econometric exploration of the consequences of such sample selection was Heckman's work in 1974. The standard example is the estimation of salary based on an analysis of working women on their own, because the decision to work involves a trade-off in which the salary an individual can potentially earn plays a role. Since this, many papers have highlighted the importance of the selection bias in human and social science surveys (Maddala, 1986). Noteworthy examples are the model for migration in the USA analysed by Nakosteen and Zimmer (1980), or that for female employment rates analysed by Mroz (1987). The most frequent use of self-selection models is for the evaluation of processing or training.

In practice, the selection bias has two sources (Heckman, 1990). It results either from respondent self-selection or a selection decision by the study managers. When mixed survey modes are used, individuals choose to belong to one group or another or only respond if the proposed medium suits them. The responses are therefore not comparable, because the sample is no longer random and the presence of respondents is determined by external factors which may also affect the variable of interest in the studied model. It is highly likely that the socio-economic characteristics and the travel behaviours of the individuals who respond using the internet are different from those of the individuals who respond to a telephone interview (Resource system group, 2002; Lozar Manfreda et al., 2008). This conclusion can be drawn as long as there is no explanatory variable missing.

The data from the Rhône-Alpes area travel survey highlights the problem of self-selection, with the non-respondents to the standard phone survey choosing whether or not to fill in the internet questionnaire. We illustrate this by proposing an explanatory model for travel. More precisely, we analyse the average number of trips per person. Our analysis shows that the internet respondents travelled less than the phone respondents (3.09 vs 3.67 daily trips). Let us consider an equation that analyses the effect of survey mode on an individual's average number of daily trips:

$$Y_i = \beta_k X_{ki} + \alpha I_i + u_i \quad (1)$$

where Y_i is the average number of trips made by respondent i (dependent variable), X_{ki} is a vector of K explanatory variables and I_i is a dummy variable that states whether the individual i responded by internet. A question which arises is whether the coefficient α measures the real impact on daily travel of the choice of responding on the internet. The answer to this question is affirmative if individuals who decide to respond on the

internet would have reported the same number of trips if they had responded in the phone situation. However, the variable I cannot be considered as exogenous in this model, as the contacted individuals chose whether or not to respond in the phone interview and whether or not fill out the online survey. Respondent self-selection must be corrected during least squares regression in order to obtain unbiased estimates of the coefficients, by using the two-stage estimation method developed by Heckman in 1979.

The central hypothesis is that random disruptions simultaneously affect the endogenous variables in the selection and mobility equations and that the error terms in the two equations may be correlated. The endogeneity of the survey mode with respect to travel arises as a result of omitted variables which are correlated to the probability of choosing the web-based survey and the number of reported trips.

5.1. Selection equation

The first stage consists of estimating the survey medium "choice" (web vs phone interview) equation using a probit model. Before beginning the econometric analysis, we select variables which make it possible to distinguish the individuals who responded on the internet. Two types of variables drew our attention in this context: the respondent's willingness to participate in the travel survey (web survey or GPS study) and some socio-demographic characteristics (gender, status, age, driving licence, whether they stated their monthly income, or declared Friday trips).

Table 3. Selection model (Probit).

Variables	Estimates	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.8e+00	1.6e-01	-11.21	< 2e16 ***
Gender_male	9.0e-02	3.1e-02	2.94	3.31e-03 **
Age	3.4e-02	6.0e-03	5.75	8.92e-09 ***
Age^2	-4.0e-04	5.7e-05	-6.98	2.96e-12***
Driving licence	3.0e-01	7.2e-02	4.24	2.23e-05 ***
Job_executive	2.3e-01	4.4e-02	5.40	6.82e-08 ***
Job_employee	-2.5e-01	4.4e-02	-5.68	1.35e-08***
No job	-9.5e-01	1.1e-01	-8.45	< 2e-16 ***
Friday reference day	5.2e-01	3.3e-02	15.79	< 2e-16 ***
Declared income	2.0e-01	4.3e-02	4.62	3.80e-06 ***
New web survey	-6.0e-01	3.9e-02	-15.53	< 2e-16 ***
GPS survey	-1.7e-01	3.6e-02	-4.64	3.50e-06 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Null deviance: 9,846 on 13,991 degrees of freedom
Residual deviance: 8,598 on 13,980 degrees of freedom

The coefficients estimated by the Probit model are significant (Table 3). An examination of their sign gives a good idea of the importance of the various factors which influence choice of the internet. Unsurprisingly, the probability of responding to the survey by internet is higher for males, executives and the middle-aged (from 30 to 60 years old). Likewise, the probability of responding online increases when individuals report their monthly income, have a driving licence and report trips made on Friday, as answers to the web survey were often done during the weekend. Conversely, the probability of using the web falls with the intention to fill in another web travel survey or to engage in a future GPS survey. After 20 minutes in front of the computer to fill in the survey, perhaps respondents do not want to participate again.

5.2. Regression model

The second stage consists of explaining the differences in travel behaviour using a specific model which includes:

- a dependent variable, the average number of trips made by individuals
- several independent or explanatory variables, the observed factors assumed to have an effect on the number of trips reported by individuals
- the inverse Mills ratio, a correction factor for each individual obtained in the first stage, which corresponds to the conditional expectation of the residues of the selection equation with respect to the survey mode
- an error term which takes account of the unobserved forces that could influence the measurement of results.

We estimate two models, one using the subsample of telephone respondents, and the other using the subsample of internet respondents. The coefficients of the explanatory variables and the selection bias variable are estimated by a least squares regression method. We only consider here the explanatory variables with a direct impact on individuals' travel (age, residential location, the number of children in the household, possession of a driving licence, the number of cars by household adults, and the job status). The variable "residential location" has five values: the larger cities (Lyon, Villeurbanne, Grenoble and Saint-Etienne) which is the reference value in the models, the PTU (transit authority perimeter) of cities listed above, the medium cities (for example: Villefranche, Vienne, Valence…), the PTU of these medium cities, and small cities. The variable "house" is related to households living in an individual house versus building housing which is the reference value. In each model, the coefficient of the selection bias variable (Mills ratio) is significant. Individuals are therefore subject to a selection bias, and this can be eliminated by using a two-stage estimation method. We thus obtain unbiased estimates. Selection bias variable converges under less restrictive assumptions than the maximum likelihood test.

Table 4. Travel model for telephone respondents (OLS).

Variables	Estimates	Std. Error	t value	Pr(>|t|)
(Intercept)	1.5e-01	2.6e-01	0.57	0.568
Age	8.7e-02	7.1e-03	12.21	< 2e-16 ***
Age^2	-1.0e-03	6.8e-05	-15.16	< 2e-16 ***
Driving licence	7.4e-01	7.6e-02	9.77	< 2e-16 ***
No. of children	6.0e-01	6.1e-02	9.93	< 2e-16 ***
No. of $children^2$	-5.1e-02	1.9e-02	-2.66	0.008 **
Cars per pers>18 years	2.7e-01	5.3e-02	5.00	5.83e-07 ***
			6.31	2.88e-10 ***
Job_executive	4.2e-01	6.7e-02		
Job_employee	3.7e-02	6.3e-02	0.57	0.567
No job	-5.6e-01	1.2e-01	-4.59	4.44e-06 ***
Residential location _3towns PTU	-2.2e-01	1.0e-01	-2.11	0.034 *
Residential location _other towns	1.5e-02	8.1e-02	0.18	0.855
Residential location _other towns PTU	-1.7e-01	8.8e-02	-1.97	0.049 *
Residential location _others	-2.1e-01	6.7e-02	-3.12	0.002 **
House	-1.7e-01	4.8e-02	-3.45	0.001 ***
Mills	5.6e-01	7.8e-02	7.21	6.12e-13 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 2.24 on 12,401 degrees of freedom
Multiple R-squared: 0.119, Adjusted R-squared: 0.118
F-statistic: 112 on 15 and 12,401 DF, p-value: < 2.2e-16

The model which is restricted to the subsample that responded to the phone survey provided interesting results, as all the variables are significant (Table 4). Living in large cities in the Rhône-Alpes area, having cars available and a

driving licence and being an executive increases the average daily number of trips. Travel also seems to be an increasing function of age and of the number of children in the household (up to a certain level). The Mills' coefficient has a positive sign: those who responded to the phone survey would have declared a lower number of trips if they had answered the web survey.

In the model which is restricted to the subsample of individuals who responded online, few of the coefficients achieved significance (Table 5). This is explained in particular by the differences in the sizes of the two samples of respondents. A total of 12,417 individuals responded to the phone survey, compared with only 1,626 on the web, a ratio of one to eight. Being an executive and having a car available have a positive impact on travel. However, the number of trips increases with age and children in the household. The Mills' coefficient has a positive sign: those who decided to answer on the web would have reported more daily trips if they had answered the phone interview.

Table 5. Travel model for internet respondents (OLS).

Variables	Estimates	Std. Error	t value	Pr(>|t|)
(Intercept)	-4.2e-01	7.9e-01	-0.54	0.590
Age	9.1e-02	2.5e-02	3.60	3.32e-04 ***
Age^2	-1.0e-03	2.4e-04	-4.29	1.93e-05 ***
Driving licence	2.2e-01	3.6e-01	0.63	0.530
No. of children	6.4e-01	1.8e-01	3.53	4.30e-04 ***
No. of $children^2$	-4.7e-03	6.3e-02	-0.07	0.941
Cars per pers>18 years	3.3e-01	1.7e-01	1.94	0.052 .
Job_executive	6.0e-01	1.7e-01	3.59	3.43e-04 ***
Job_employee	1.2e-01	1.8e-01	0.69	0.489
No job	-6.8e-01	5.6e-01	-1.22	0.221
Residential location _3towns PTU	-2.6e-01	3.2e-01	-0.82	0.414
Residential location _other towns	4.1e-01	2.4e-01	1.70	0.090 .
Residential location _other towns PTU	2.3e-01	2.6e-01	-0.88	0.382
Residential location _others	-2.5e-01	1.9e-01	-1.30	0.196
House	-9.2e-02	1.4e-01	-0.64	0.520
Mills	6.2e-01	2.0e-01	3.18	1.50e-03 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 2.32 on 1,559 degrees of freedom
Multiple R-squared: 0.12, Adjusted R-squared: 0.11
F-statistic: 13.8 on 15 and 1,559 DF, p-value: < 2.2e-16

5.3. Model stability test

The purpose of this is to test whether the coefficients of the two-stage model are stable and to quantify any impact of the survey mode. The results from the regression model applied to the explanatory variables and main interactions are set out in Table 6. Only one variable appears to interact significantly with the response mode. This is the number of children in the household. However, the Mills ratio coefficients for the two samples do not differ significantly.

In our analyses we retain those interactions that are significant at the 10% error level, with the exception of the coefficient of the survey mode variable (p-value = 15.4%).

The data collection mode (web or telephone) had a direct impact on travel levels. The number of trips fell by 0.6 when subjects answered the questionnaire online.

Most of the internet respondents are employed as executives and office workers with jobs in the city centre. The further their home is from the city centre, the greater the distance between their home and their work. This reduces the time available for activities that are less constrained which reduces their levels of travel.

The mobility of respondents increased with the presence of children. They have to manage family activities in addition to their work. This effect is amplified in the case of internet respondents. Perhaps this is because their more demanding jobs (a large number of them are executives) give them less scope for making trips, as two-thirds of the internet respondents are economically active.

The coefficient of the variable No. of children*mode is thus able to quantify the impact of the survey mode on travel behaviour. The mode impact on the number of children was 0.54 per person for phone respondents vs 0.73 for internet respondents (for one child). It is likely that as the number of children in a household increases so do the number of trips made by its members (for shopping or leisure purposes, for example).

Table 6. Stable model (OLS).

Variables	Estimates	Std. Error	Estimates	Std. Error
(Intercept)	1.4e-01	2.5e-01	0.54	0.588
Age	8.7e-02	6.9e-03	12.67	< 2e-16 ***
Age^2	-1.0e-02	6.5e-05	15.80	< 2e-16 ***
Driving licence	7.4e-01	7.5e-02	9.82	< 2e-16 ***
No. of children	5.9e-01	5.9e-02	10.07	< 2e-16 ***
No. of $children^2$	-4.7e-02	1.8e-02	-2.56	1.04e-02 *
Cars per pers>18 years	2.7e-01	5.1e-02	5.32	1.07e-07 ***
Job_executive	4.5e-01	6.2e-02	7.29	3.38e-13 ***
Job_employee	5.0e-02	6.0e-02	0.84	0.403
No job	-5.6e-01	1.2e-01	-4.81	1.52e-06 ***
Residential location _3towns PTU	-2.2e-01	9.8e-02	-2.29	0.022 *
Residential location _other towns	-2.8e-02	7.7e-02	-0.37	0.712
Residential location _other towns PTU	-1.8e-01	8.3e-02	-2.16	0.031 *
Residential location _others	-2.1e-01	6.3e-02	-3.35	8.0e-04 ***
	-1.6e-01	4.5e-02	-3.49	4.90e-04 ***
House				
Mills	5.7e-01	7.2e-02	7.90	3.11e-15 ***
Survey mode	-5.8e-01	4.1e-01	1.43	0.154
Driving licence: Mode	-3.9e-01	3.2e-01	1.25	0.213
No. of children * Mode	1.9e-01	7.1e-02	2.61	0.009**
Age * Mode	5.2e-03	4.4e-03	1.18	0.237

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 2.25 on 13,975 degrees of freedom
Multiple R-squared: 0.121, Adjusted R-squared: 0.120
F-statistic: 120 on 16 and 13,975 DF, p-value: < 2.2e-16

6. Results mainly in line with the previous Lyon survey

A project of conducting a web survey in parallel with a household travel survey was conducted in Lyon in 2006. In the main survey, households were interviewed at home face-to-face: 11,000 interviews were conducted between October 2005 and March 2006. Households who refused to respond (about 31%) or were not contactable after a certain number of attempts (about 16%) received a letter to invite them to respond by the web. At the end, 369 households gave a complete and usable answer, which represented an answer rate of 8.5%, very similar to the Rhône-Alpes survey as there were two reminders for the Lyon web survey and only one for only a small part of the sample for the Rhône-Alpes web survey. Differences in socio-economic characteristics of respondents stand out clearly between the face-to-face and web surveys. The households who answered the web survey had more members (2.60 persons vs 2.37), were more motorised and better equipped in communication tools and had a higher annual income by unit of consumption (20,000 euros vs 15,000 euros on average). This difference could be explained by

the low penetration rate of internet in 2006 (which concerned first high income households). Moreover, those who have a connection at the workplace have generally superior occupations, with a higher level of income.

There was no difference in terms of respondent's gender, but the proportion of active persons (between 30 and 59 years) who answered on the web was higher than in the standard survey. They belonged to high social and occupational groups (39% were executives or had a superior intellectual profession vs 14%) and had a high level of studies (about two-thirds made studies vs only 30% in the standard survey). Web respondents more frequently had a driving licence and worked or studied more often in the town centre (27% vs 16%), where there are head offices of companies. Socio-demographic differences between web and telephone on the one hand and face-to-face and web on the other hand appear quite similar which might be interpreted as a difference between web and surveyed survey mode.

Web respondents declared fewer daily trips than the respondents interviewed face-to-face (3.63 vs 2.97). This gap was explained by a higher number of immobiles (19% vs 11%) and a lower mobility of mobiles in the web survey (3.71 vs 4.08 daily trips). They most often drive a car (58% vs 44%), and walk less (15% vs 29%). The purpose at destination "Work" was over-represented in the web survey, compared to the standard survey (23% vs 15%), because three-quarters of the web respondents were active persons. Web respondents covered longer distances on average (4.61 km vs 3.78 km), and their trips have a longer duration (25.54 minutes vs 19.25 minutes).

We identified a selection bias, and showed that data collection mode (web or face-to-face) had a direct impact on mobility. The number of trips fell by 0.5 when subjects answered the questionnaire online.

These results are in line with the new web survey conducted in the Lyon area even if the perimeter is larger and the main survey mode is not the same (telephone vs face-to-face). Internet penetration rate has increased and most of the inhabitants could use the web at home or at work. It reduces the differences a little in terms of the respondent profile between web and telephone survey, but the web sample is still specific compared to the face-to-face or telephone sample. Using the margin calibration method, it is possible to significantly reduce differences. On the other hand, some gaps still remain on mobility patterns. The number of daily trips declared by web respondents is lower than trips given by phone to an interviewer and differences are again similar for the two experiments. We can suppose that the questionnaire conducted by an interviewer allowed more smaller trips to be collected.

7. Conclusion

Transport constitutes a strategic tool for urban policies. Knowing with accuracy the travel pattern of inhabitants is necessary to ensure sustainable development and management of infrastructure. The relevance of selected investments will depend on the quality of gathered data. It thus becomes crucial to identify a reliable survey protocol to obtain representative data at reasonable cost. In a perfect world, the methodology implemented would make it possible to survey the entire selected population according to a good quality/cost ratio, to reduce the non-response rate, to respect the cultural differences and to develop a reliable forecasting model (Morris, 2003). Unfortunately, it is not possible to reach all these objectives in a household travel survey. Recent technological advances, like the expansion of the internet, open new windows which have to be studied. To improve data quality, the web is used in the Rhône-Alpes area to support the telephone method. The aim of the experiment was to capture some households who did not want to respond to the traditional phone survey. The web is not considered as an alternative media to fill out the questionnaire, but as a way to survey non-respondents to traditional survey media. This method can increase survey quality by generating information usually ignored. It could be a way to balance the growth of refusals endured by interviewers over many years.

This analysis is rather complex, as we have to take into account socio-economic differences between individuals who answered on the web and those who completed the phone interview, differences of mobility patterns between these two samples and finally differences which can be imputed to the media only. Data shows that internet respondents declare fewer trips than respondents to the phone survey. This gap can be explained by a double effect: a higher immobility of the web respondents and a lower statement of their daily trips. An analysis by mode or purpose shows that the differences mainly concern a lower proportion of walking, and a sharply weaker participation in the activities of shopping-services. The data is consistent with the hypothesis of an under declaration

attributable to the web, as far as we know, that the risk of omitting trips concerns especially short trips in terms of distance and duration and less constrained trip purposes.

Attention has been paid to data comparability issues in certain disciplines, but it has not yet received enough attention in the field of transport. We propose that the two-stage model proposed by Heckman in 1974 should be adapted. Application to the Rhône-Alpes Area household travel survey in which two survey media were used has revealed the presence of a selection bias. The traditional estimation of a travel equation using the ordinary least squares method for the entire population therefore ignores an important factor: the determinants of survey mode choice. On the other hand, taking account of this selection bias provides a means of estimating unbiased relationships. It would be beneficial to develop this type of approach in the area of travel surveys in order, ultimately, to propose methods for combining samples obtained with different methodologies in order to increase the statistical representativeness of the surveys and reduce the non-response problem.

The two web surveys in Lyon area and in Rhône-Alpes Region allow some conclusions to be made regarding web survey bias in comparison to face-to-face and telephone surveys where an interviewer is present to make some reminders if necessary to interviewees. Firstly even if the web penetration rate has increased (50% in 2006 and 83% in 2013 for France, http://www.internetworldstats.com/stats4.htm), the main socio-demographic biases are still present even if slightly reduced. Web respondents still have higher income, educational level, and occupation group. They also belong to households with more cars and individuals. Older persons are less present. It also the case for younger persons but it is probably due to the contact method. When the socio-demographic biases are corrected by a weighting procedure, again we observe similar differences between face-to-face and web surveys on the one hand and telephone and web surveys on the other hand. Again the differences are slightly reduced in 2013 compared to the 2006 surveys. But we still observe a lower declaration of the number of trips, mainly due to a higher level of immobility in the web survey. When comparing mobility only for those who are mobile we obtain nearly 0.5 trips less in the web survey vs face-to-face or telephone surveys. Differences are mainly concentrated on walking trips and for less constrained purposes which could explain under declaration due to the absence of an interviewer. In contrast, average distance and duration of web trips are higher and web respondents return home later than face-to-face and telephone respondents. Lastly we observe similar response rates if we consider that the number of reminders are different for the two web surveys (around 10%).

The feasibility of web household travel surveys has now been demonstrated with several examples around the world. This research thus represents an important step for modelling and surveys. But web respondents cannot be regarded today as representative of the French population. We therefore still recommend combining survey modes in order to increase response rates. A direction for further research is to allocate “correcting coefficients” to the internet respondents with a view to integrating this sample with the phone or face-to-face respondents in order to correct the entire resulting dataset to make it representative of the whole population.

Acknowledgement

The authors would like to thank Rhône-Alpes Region for giving them access to the 2013-2015 regional travel survey data collected in Rhône-Alpes (Enquête déplacements région Rhône-Alpes 2013-2015, Conseil régional, vague 1). However, the authors alone are responsible for the contents of this paper. For this article, data has not been weighted to represent the whole regional population. Therefore it cannot be used to represent Rhône-Alpes mobility.

References

Alsnih, R., 2006. Characteristics of Web-based surveys and applications in travel research, in “*Travel Survey Methods, Quality and Future Directions*”. In: Stopher, P.R., Stecher, C. (Ed.). Elsevier, pp. 569-592.

Alsnih, R., 2004. New technology and travel surveys: the way forward, Working Paper ITS-WP-04-11. Institute of Transport Studies, University of Sydney, April 2004, Australia.

Atrostic, B.K., Bates, N., Burt, G., Silberstein A. 2001. Nonresponse in US Government Household Surveys: Consistent Measures, Recent Trends and New Insights. Journal of Official Statistics 17, 209-226.

Bayart, C., Bonnel, P., 2012. Combining Web and Face-to-face in Travel Surveys: Comparability Challenges. Transportation 39, 6, 1147-1171.

Bayart, C., Bonnel, P., Morency, C., 2009. Survey mode integration and data fusion: methods and challenges, in “*Transport Survey Methods: Keeping up with a Changing World*”. In: Bonnel, P., Lee Gosselin, M., Zmud, J., Madre J.-L. (Ed.). Emerald Press, pp: 587-611.

Bethlehem, J., 2010. Selection Bias in Web Surveys. International Statistical Review 78, 161-188.
Bonnel, P., 2003. Postal, telephone and face-to-face surveys: how comparable are they? in "*Transport Survey Quality and Innovation*". In: Stopher, P.R., Jones, P.M. (Ed.). Pergamon, Oxford, pp. 215-237.
Bourbonnais, P.L., Morency, C., 2012. Web-based travel survey: a demo, in "*Transport Survey Methods: Best Practice for Decision Making*". In: Zmud, J., Lee-Gosselin, M., Munizaga, M., & Carrasco, J.A. (Ed.). Emerald, Bingley, pp. 207-223.
CERTU 1998. L'enquête ménage déplacements « méthode standard », Collections du CERTU, éditions du CERTU, Lyon, 295p.
Christensen, L., 2012. The Role of Web Interviews as Part of a National Travel Survey, in: "*Transport Survey Methods: Best Practice for Decision Making*". In: Zmud, J., Lee-Gosselin, M., Munizaga, M., & Carrasco, J.A. (Ed.). Emerald, Bingley, pp. 115-153.
Couper, M.P., Bosnjak, M., 2010. Internet surveys, in "*The Handbook of Survey Research*". In: Marsden, P.V., & Wright, J.D. (Ed). Bingley, Emerald, UK, pp. 527-556.
Couper, M.P., 2008. Designing effective web surveys, New York: Cambridge University Press.
Couper, M.P., Traugott, M.W., Lamias, M.J., 2001. Web Survey Design and Administration. Public Opinion Quarterly 64, 4, 464-494.
Curtin, R., Presser, S., Singer, E., 2005. Changes in Telephone Survey Nonresponse over the Past Quarter Century. Public Opinion Quarterly, 69, 87-98.
Dillman, D.A., 2007. Mail and Internet surveys: The tailored design method, 2nd edition. New York: John Wiley.
Dillman, D.A., Bowker, D.K., 2001. The Web Questionnaire Challenge to Survey Methodologists, in: "*Dimensions of Internet Science*". In: Reips, U.D., Bosnjak, M. (Ed.). Pabst Science Publishers, Lengerich, Germany.
Dillman, A., Phelps, G., Tortora, R., Swift, K., Kohrell, J., Berck, J., Messer, B.L., 2009. Response Rate and Measurement Differences in Mixed-mode Surveys Using Mail, Telephone, Interactive Voice Response (IVR) and the Internet. Social Science Research, 38,1–18.
Dillman, D.A., Tortora, R.D., Bowker, D.K., 1998. Principles for Constructing Web Surveys, SESRC Technical Report 98-50, Pullman, Washington.
De Leeuw, E.D., 2005. To Mix or Not to Mix Data Collection Modes in Surveys. Journal of Official Statistics 21, 233-255.
Groves, R.M., Lyberg, L., 2010. Total Survey Error: Past, Present and Future. Public Opinion Quarterly 74, 849-879.
Gunn, H., 2002. Web-based Surveys: Changing the Survey Process. First Monday 7,12, December 2002.
Heckman, J., 1979. Sample Selection Bias as a Specification Error. Econometrica, 47,1, 153-161.
Heckman, J., 1990. Varieties of Selection Bias. The American Economic Review, 80, 2. Papers and Proceedings of the 102nddecl Annual Meeting of the American Economic Association, 313-318.
Kagerbauer, M., Manz, W., Zumkeller, D., 2012. Analysis of PAPI, CATI and CAWI Methods for a Multi-Day Household Travel Survey, in: "*Transport Survey Methods: Best Practice for Decision Making*". In: Zmud, J., Lee-Gosselin, M., Munizaga, M., Carrasco, J.A. (Ed.). Emerald, Bingley, pp. 289-303.
Lozar Manfreda, K., Bosnjak, M., Berzelak, J., Haas, I., Vehovar, V., 2008. Web Surveys versus Other Survey Modes. International Journal of Market Research, 50, 79–104.
Maddala, G.S., 1986. Limited dependant and qualitative variables in econometrics. Cambridge University Press, London.
http://firstmonday.org/issues/issue7_12/gunn/index.html
Mokhtarian, P.L., Chen, C., 2004. TTB or not TTB, that is the Question: A Review and Analysis of the Empirical Literature on Travel Time (and Money) Budgets. Transportation Research Part A: Policy and Practice, 38, 9:10, 615-689.
Morris, J., Adler, T., 2003. Mixed mode survey, in "*Transport Survey Quality and Innovation*". In: Stopher, P.R., Jones, P.M. (Ed.). Pergamon, Oxford, pp. 239-252.
Mroz, T.A., 1987. The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions. Econometrica, 55, 4, 765-799.
Nakosteen, R.A., Zimmer, M.A., 1980. Migration and Income: the Question of Self-selection. Southern Economic Journal, 46, 840-851.
Richardson, A.J., 2000. Behavioural Mechanisms of Non-Response in Mailback Travel Surveys, 79th Transportation Research Board, Washington, 18p.
Roy, A.D., 1951. Some Thoughts on the Distribution of Earnings, Oxford Economic Papers 3, 135–146.
Stopher, P.R., Jones, P.M., 2003. Summary and future directions, in "*Transport Survey Quality and Innovation*". In: Stopher, P.R., Jones, P.M. (Ed.). Pergamon, Oxford, pp. 635-646.
Thériault, M., Lee-Gosselin, M., Alexandre, L., Théberge, F., Dieumegarde, L., 2012. Web versus Pencil and Paper Surveys of Weekly Mobility: Conviviality, Technical and Privacy Issues, in "*Transport Survey Methods: Best Practice for Decision Making*". Zmud, J., Lee-Gosselin, M., Munizaga, M., Carrasco, J.A. (Ed.). Emerald, Bingley, pp. 225-245.
Tourangeau, R., Conrad, F.G., Couper, M.P., 2013. The science of web surveys, Oxford University Press, 198p.
Yun, G.M., Trumbo, C.W., 2000. Comparative Response to a Survey Executed by Post, E-mail & Web Form. Journal of Computer-Mediated Communication, 6,1. http://jcmc.indian